

## Anmerkungen zu Birger Hjørland und Jeppe Nicolaisen: Bradford's Law of Scattering: Ambiguities in the Concept of "Subject"

von Walther Umstätter

Es hat zahlreiche Versuche gegeben, *Bradford's Law of Scattering* (BLS) zu korrigieren, zu präzisieren oder auch zu falsifizieren. Diese Versuche sind teilweise zum Scheitern verurteilt, wenn man das BLS als das nimmt, was es ist, eine Regelhaftigkeit, die die alte 80:20 Regel der Bibliotheken (etwa 80% der Nachfrage können mit etwa 20% des Bestandes abgedeckt werden) auf die dokumentarischen Belange der Zeitschriftennutzung erweiterte. Garfield hat dieses Postulat für seinen *Science Citation Index* (SCI) als „*Garfield's Law of Concentration*“ erfolgreich eingesetzt, als er nach eigenen Angaben herausfand, dass 25 Zeitschriften im SCI etwa 20-25% der 4 Mio. Zitationen 1969 abdeckten<sup>1</sup>. Bemerkenswerterweise fehlt dieses Zitat in der Arbeit von Hjørland und Nicolaisen.

Die Grundaussage im BLS war, dass sich eine immer gleiche Menge an Literatur auf verschiedene, thematisch geordnete Zeitschriften wie  $n^0:n^1:n^2:n^3 \dots$  verteilt. Schon an dieser Schreibweise ist erkennbar, dass es sich lediglich um eine Präzisierung der 80:20 Regel handelt. Wenn wir somit in einer thematisch zentralen Zeitschrift z.B. pro Jahr 50 für uns relevante Aufsätze finden (z.B. „Origin of Life“<sup>2</sup>), und etwa 30 Kernzeitschriften brauchen um 100 zu finden, sowie etwa 27.000 um 200 zu finden, so wäre das BLS erfüllt.

Bei jeder natürlichen Beobachtung in der Fachliteratur würde ein Streit darüber, ob man beispielsweise mit 1:30:900:27000 oder 1:31:931:27.931 rechnen müsste, schon allein durch die auftretenden Streuungen ad absurdum geführt. Hier gilt wie so oft: „Der Mangel an mathematischer Bildung gibt sich durch nichts so auffallend zu erkennen, wie durch maßlose Schärfe im Zahlenrechnen.“<sup>3</sup>

Es gibt bekanntlich schon in Bradfords Buch leicht abgewandelte Angaben zu seiner Beobachtung. Insbesondere im Kernbereich wurde über den *Groos droop* diskutiert. Solche Detaildiskussionen sind aber angesichts realer Erfahrungen auf diesem Gebiet abwegig, denn sie sind bekanntlich stark abhängig von der Recherchestrategie und der Relevanzbewertung.

Das BLS lässt sich in der Formel  $p=k^{((a/c)-1)}$  fassen<sup>4</sup>, wobei  $p$ =Anzahl der zu untersuchenden Periodika,  $a$ =Anzahl der Aufsätze in diesen Periodika,  $c$ =Anzahl der Aufsätze im 1. Periodikum (Zentralzeitschrift) und  $k$ =Anzahl der Kernzeitschriften ist, die wieder  $c$  Aufsätze enthalten. Damit sind  $c$  und  $k$  Konstanten, die die Funktion bestimmen. Diese beiden Konstanten sind ein Ausdruck dafür, ob sich ein Thema (concept, problem, subject, word,

<sup>1</sup> Garfield, E. "The Mystery of the Transposed Journal Lists -- Wherein Bradford's Law of Scattering is generalized. According to Garfield's Law of Concentration," *Essays of an Information Scientist*, Vol:1, p.222, 1962-73. <http://garfield.library.upenn.edu/essays/v1p222y1962-73.pdf>

<sup>2</sup> Umstätter, W. und Rehm, M.: Einführung in die Literaturdokumentation und Informationsvermittlung. K.G. Saur Verl. München (1981)

<sup>3</sup> nach C.F. Gauss

<sup>4</sup> Umstätter, W. und Rehm, M.: Einführung in die Literaturdokumentation und Informationsvermittlung. K.G. Saur Verl. München (1981)

etc.<sup>5)</sup> mehr oder weniger flach über die Zeitschriftentitel verteilt. Damit wird der eigentliche Gegenstand der Betrachtung bei Hjørland and Nicolaisen gegenstandslos, weil die Funktion entsprechend anpassungsfähig ist. Dies ist auch der Grund, warum Garfield das *BLS* für „single disciplines“ auch auf „science as a whole“ ausweiten konnte. Umgeformt ergibt sich  $a=c((\ln p/\ln k) + 1)$ . Für eine Spezialbibliothek mit 900 laufenden Zeitschriften (Beispiel s. „Origin of Life“) lässt sich ausrechnen, dass in 20%=180 Zeitschriften  $126=50*(\ln 180/\ln 30) + 1$  Aufsätze, d.h. 126/150 (150 Aufsätze in den 900 Zeitschriften) =84% abgedeckt werden könnten. Mit nur 30 Zeitschriften würden 20%=6 Zeitschriften 76 Prozent abdecken.<sup>6</sup>

Verallgemeinert man diese Überlegung dahingehend, dass man für immer breitere bzw. engere Themen entsprechende Verteilungen nach dem BLS konstruiert, so ergibt sich<sup>7</sup> für beispielsweise 10.000 Publikationen, dass für die „Zentralzeitschrift“ nicht eine, sondern 100 Zeitschriften (für  $n^0$ ) anzunehmen sind, und dass entsprechend alle weiteren folgenden Ränge durch 100 geteilt werden müssen.

Vergleicht man damit idealtypisch die Abnahme der Publikationen pro Zeitschriftentitel, im Sinne des BLS über einen weiten Bereich von etwa 8 bis 200.000 Publikationen pro Jahr und Thema (Abb.1 gestrichelte Linien), so zeigt sich, dass bei etwa 200 Publikationen pro Jahr (hervorgehobenen Kurve) eine neue Zeitschrift entsteht. Dies entspricht dem beobachteten Fall, der bereits 1984<sup>8</sup> bei der Entstehung der Zeitschrift „Origin of Life“ beschrieben wurde. Wir können somit zusammenfassen, dass ein neues Spezialgebiet bei etwa 200–300 Publikationen pro Jahr beginnt, wenn sich damit auch eine neue Zentralzeitschrift bildet.

Unter dieser Betrachtung haben wir es heute mit ~100.000 Zeitschriften und ebensoviel Spezialgebieten, die von 10 Mio. Wissenschaftlern betreut werden, zu tun. Wenn rein rechnerisch jede wissenschaftliche Person eine Zentralzeitschrift und rund 100 Kernzeitschriften regelmäßig sichtet, dann werden diese im Durchschnitt von 10.000 Wissenschaftlern weltweit gesichtet. In der Realität folgt aber diese Nutzung auch einer schiefen Verteilung, weil auch jeder einzelne Wissenschaftler unbewusst, aber aus eigener Erfahrung dem BLS folgt. Während also beispielsweise „Nature“ oder „Science“ von fast allen Naturwissenschaftlern frequentiert werden, gibt es zahlreiche Zeitschriften, die nur wenige Leser haben.

Die Abbildung 1 macht auch deutlich, dass das BLS weniger dem Potenzgesetz (*power law*) folgt, wie es seit 1960<sup>9</sup> immer wieder behauptet wird, sondern einer logarithmischen Funktion, da sich die Geraden aus einer halblogarithmischen Darstellung und nicht aus einer Logarithmierung von x- und y-Achse ergeben. Dies ist aber eine Frage der Rangbildung.

<sup>5</sup> Hjørland, B. and Nicolaisen, J.: Bradford's Law of Scattering: Ambiguities in the Concept of "Subject". 2005. S.103

<sup>6</sup> vgl. Umstätter, W.: Szientometrische Verfahren. In: Grundlagen der praktischen Information und Dokumentation / Rainer Kühlen u.a. (Hrsg.) 5., vollst. neu gefasste Aufl. München: Saur, 2004. S. 237-243

<sup>7</sup> Umstätter, W.: Was ist und was kann eine wissenschaftliche Zeitschrift heute und morgen leisten. [http://www.ib.hu-berlin.de/~wumsta/infopub/pub2001f/JB02\\_143-166.pdf](http://www.ib.hu-berlin.de/~wumsta/infopub/pub2001f/JB02_143-166.pdf)

<sup>8</sup> Umstätter, W. und Rehm, M.: Bibliothek und Evolution. Nachr. f. Dok., 35, 1984. S.237-249

<sup>9</sup> Kendall, M.: The bibliography of operational research. Oper Res. Quart. (1/2), 1960. S. 31-36

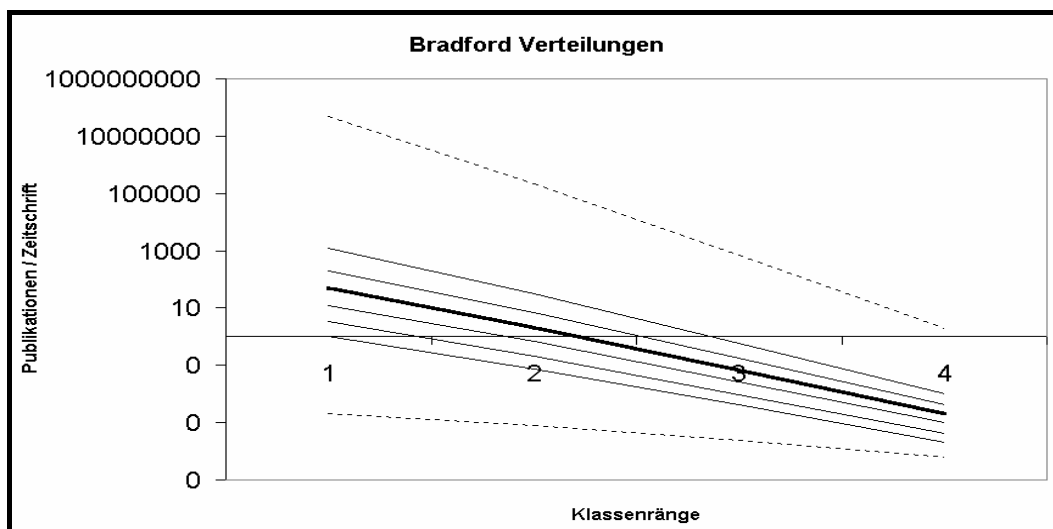


Abb. 1: Vergleich der idealtypischen Abnahme der Publikationen pro Zeitschriftentitel im Sinne des *Bradford's Law of Scattering* über einen weiten Bereich von etwa 8 bis 200.000 Publikationen pro Jahr (gestrichelte Linien). Bei etwa 200 Publikationen/J (hervorgehobenen Linie) entsteht eine neue Zeitschrift.

Im Prinzip ist das BLS damit eine funktionale Darstellung der Interdisziplinarität bzw. der „unity of science“<sup>10</sup>, das durch die beiden Konstanten  $c$  und  $k$ , der Größe und der Vernetzung eines Themas in der Gesamtwissenschaft Rechnung trägt. Wenn in diesem Zusammenhang Hjørland und Nicolaisen meinen, dass „the social sciences are generally considered very interdisciplinary“, und dass BLS dem nicht genügend Rechnung trägt, so ist dies zweifellos eine Fehleinschätzung des BLS, weil dieses nicht a priori eine endgültige Festlegung trifft, sondern gerade durch  $c$  und  $k$  an die jeweilige Thematik, Problematik, Fragestellung, Recherche, etc. angepasst wird.

Es war zweifellos ein großer Erfolg der Bibliotheks- bzw. Dokumentationswissenschaft eine Formel für die Interdisziplinarität aller Wissenschaftsbereiche gefunden zu haben. Um so bedauerlicher ist es, dass die eigene Fachwelt über diese grundlegende Bedeutung des BLS so wenig Kenntnis genommen hat, wie es in diesem Aufsatz erneut deutlich wird. Ansonsten sollte man nicht übersehen, wie interdisziplinär

- Bibliothekswissenschaft ist, da es Bibliotheken in allen Fachgebieten gibt;
- Mathematik ist, da sie in fast allen Fachgebieten nutzbar ist;
- Jura ist, da es kaum eine Disziplin im rechtsfreien Raum gibt;
- Physik, Philosophie, Linguistik, Kunst, Chemie, Biologie, Astronomie sind, ...

Die Nutzung und das Verständnis des BLS ist keine Frage eines „naïve realism“<sup>11</sup> sondern, weil die Frage nach der Semiotik eines zu betrachtenden Gegenstandes, eines Wortes oder auch einer komplexen Problematik auf einer weit höheren Ebene liegt, eine Frage der wissenschaftlichen Konzentration von Themen (*Garfield's Law of Concentration*). Sie beruht auf der Wahrscheinlichkeit, dass Wissenschaftler bei gemeinsamen Problemlösungen bestimmte Zeitschriften bevorzugen. Das BLS ist so gesehen ein Maß für die *collaboration* in der Wissenschaft. Es hat insofern mit „axiological arguments“ nur sehr indirekt zu tun.

<sup>10</sup> Hjørland, B. and Nicolaisen, J., 2005. S.100

<sup>11</sup> vgl. ebd., S.101

Wenn beispielsweise festgestellt wird, „Bradford multipliers vary from zone to zone“<sup>12</sup>, so müssen zwei Fragen unterschieden werden:

1. Ist die inhaltliche Festlegung beim „subject“ konstant, oder variiert sie von Zone zu Zone?
2. Ist die Variation ein systematischer Fehler oder eine zufällige Streuung?

Grundsätzlich geht es bei jeder Verteilung im Sinne Bradfords um inhaltliche Fragen, wie „Applied Geophysics“ oder „Lubrication“. Insofern ist es eine Fehleinschätzung, wenn es beim „lexical scattering“ heißt, es müsse dem „Zip’s law“ folgen, weil sozusagen Worte gesucht werden. Das Zipfsche Gesetz gilt aber nur in einem Text mit der entsprechenden Syntax und dem darin enthaltenen *principle of least effort*. Beim BLS geht es aber grundsätzlich um die Verteilung von Begrifflichkeiten auf Dokumente. Darin liegt auch die bereits erwähnte Fehleinschätzung, dass das BLS ein Potenzgesetz sei.

Am Schluss sei noch darauf hingewiesen, dass die Aussage „nobody have thus far tried to outline the consequences of different conceptions of 'subjects' for Bradford’s law“<sup>13</sup> die wichtige Diskussion vieler Rechercheure über Relevanz, Precision, Recall, Noise und Pertinenz völlig vernachlässigt. Bei diesen „Onlinern“ wird das „Bradfordizing“ vermutlich weitaus häufiger eingesetzt, als Hjørland und Nicolaisen vermuten. Denn es erleichtert nicht nur das Auffinden von Aufsätzen und hilft Anfängern rasch zu erkennen, welche Zeitschriften für sie besonders wichtig sind, es erhöht auch die Wahrscheinlichkeit, gerade in diesen Quellen noch weitere wichtige Publikationen zu finden, was dann allerdings den *Groos droop* durchaus beeinflussen kann.

---

<sup>12</sup> vgl. ebd., S. 96

<sup>13</sup> vgl. ebd., S. 97